2D MAPPING OF CELL NUCLEI BASED ON CONTOUR WARPING

Jonas De Vylder[1,*] *, Winnok De Vos*[2]*, Wilfried Philips*[1]

[1]Department of Telecommunications and Information Processing, IBBT - Image Processing and Interpretation Group, Ghent University, St-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
[2]Department of Molecular Biotechnology, Bio-imaging and Cytometry Unit, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium

ABSTRACT

The dynamics of genome regions are associated to the functional or dysfunctional behaviour of the human cell. In order to study these dynamics it is necessary to remove all perturbations coming from movement and deformation of the nucleus, i.e. the container holding the genome. In literature models have been proposed to cope with the transformations corresponding to nuclear dynamics of healthy cells. However for pathological cells, the nucleus deforms in an apparently random way, making the use of such models a non trivial task. In this paper we propose a mapping of the cell nucleus which is based on the matching of the nuclear contours. The proposed method does not put constraints on the possible shapes nor on the possible deformations, making this method suited for the analysis of pathological nuclei.

1. INTRODUCTION

The internal organisation of the human cell nucleus in space and time is essential for its function. Within the limited space of the nucleus the entire genome as well as many proteins are accommodated in a non-random manner. Of particular interest are telomeres, the ends of chromosomes, which show a spatiotemporal behaviour that is functionally relevant to the cell and organism. Telomeres are arranged in distinct patterns and display mobility regimes at different time scales. Diverse biological processes such as telomere maintenance, cell ageing and apoptosis, i.e. cell death, are associated with specific telomere movements [11]. Likewise, altered telomere dynamics and by expansion nuclear protein dynamics are associated with specific diseases such as laminopathies and cancer [8, 4]. Therefore, quantitative studies of nuclear dynamics may help revealing novel mechanisms of dysfunction and disease.

However, analyses of dynamics in time-lapse microscopic image data sets are hampered by global cell motion and deformation. This superimposes a motion on the submicron dynamics that needs to be removed. Gladilin et al. have proposed a mapping based on a spherical model [5]. This model can cover for most of the cellular and nuclear displacements (translation, rotation and small affine transformations) that occur in normal cells during interphase. However, during cell division and in certain pathological conditions, such as laminopathies, the nuclear shape alters dramatically [4]. In Fig. 1.a an example of such a pathological nucleus is shown. As can be seen is the spherical model no longer valid. Hence a method is required that allows reliable motion measurements of subnuclear features in a deformable volume without imposing any shape constraints.

In this paper we propose a new mapping technique which maps two nuclei between consecutive time points without imposing any constraint such as a spherical shape. Instead we will extract the contours out of segmented micrographs. These contours will be matched in such a way that we retrieve a point to point correspondence. Based on this matching a mapping of the full nucleus is calculated using polyharmonic splines. This paper is arranged as follows: in the next section we describe a simple segmentation technique, which can be used to extract the nucleus out of micrograph. Section 3 provides a detailed description of the contour matching algorithm. The interpolation using polyharmonic splines is described in section 4. In section 5 the validation results are explained and discussed. Section 6 recapitulates and concludes.

2. SEGMENTATION

Since the proposed method estimates the topological changes based on the deformation of the nucleus, we first need to segment the nucleus. In this paper we start from fluorescent 3D micrographs captured by a confocal microscope. We project the 3D micrograph on a 2D image. This is done by taking the average in the z direction, i.e.

$$I(x,y) = \frac{1}{n}\sum_{z=1}^{n} M(x,y,z) \tag{1}$$

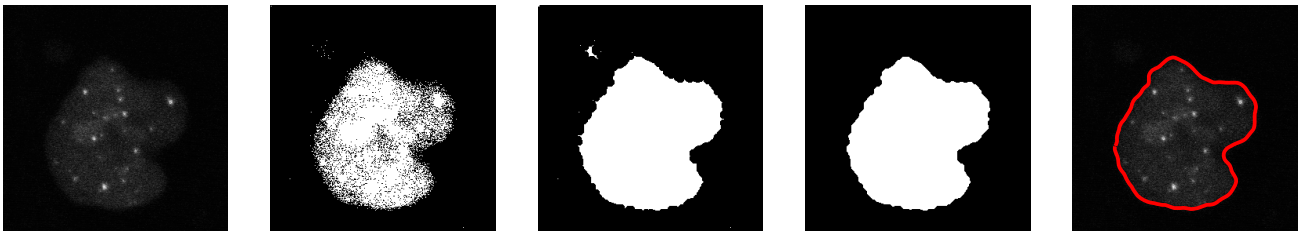
where I is a 2D image, M is the 3D micrograph and n is the z-dimension of M. An example of such a z-stack projection is shown in Fig. 1.a. Depending on the microscope system it is of course possible to start immediately from 2D micrographs such as e.g. widefield images.

The 2D image is now thresholded using Otsu's thresholding [10]. This method calculates the threshold in order to minimize the variance of the foreground, σ_f, and the variance of the background, σ_b, i.e.

$$t = \arg\min_t \left(p_f(t)\sigma_f + p_b(t)\sigma_b\right) \tag{2}$$

where $p_f(t)$ is the probability that a pixel in the image is higher than t, while $p_b(t)$ calculates the probability the pixel is lower than t. The result of Otsu thresholding can be seen in Fig. 1.b. Due to noise, not all desired pixels are considered to be foreground, e.g. part of the nucleus, whereas some background pixels are considered to be foreground. Missing foreground pixels are generally in the vicinity of detected foreground pixels. By using a morphological closing these

* Corresponding author: phone: +32 9 264 3416, email: jonas.devylder@telin.ugent.be



(a) (b) (c) (d) (e)

Figure 1: An example of the different steps in the segmentation algorithm. (a) shows the projection of the 3D micrograph on a 2D image. (b) demonstrates the result of Otsu thresholding. (c) and (d) are respectively the morphological opening and closing. (d) shows the filtered contour plotted on top of the image.

missing pixels are added to the foreground. The falsely detected background pixels are discarded using a morphological opening, as can be seen in Fig 1.c-d. Finally the border delineating the segment is filtered with a Gaussian kernel,i.e. we assume the nucleus has a smooth surface. The final result of the segmentation can be seen in Fig. 1.e.

This segmentation algorithm is developed for a specific dataset which consists of isolated cell nuclei which have good contrast compared to the background. This is off course dependent on the application and the dyes used for the fluorescent micrographs. The method for nuclei mapping proposed in this paper can also be used with more complex images, but would then acquire a more suitable segmentation technique such as described in [7, 9, 12, 2].

3. CONTOUR WARPING

In this section contours delineating the nucleus in subsequent frames will be matched, i.e. points belonging to the contour in a frame will be linked to their counterpoints on the contour in the next frame. In order to do so, we calculate a signature out of the contour, then based on this signature the matching is done.

3.1 Contour Signatures

Consider the contour $\mathbf{c}(t) = (\mathbf{x}(t), \mathbf{y}(t))$ with $t \in \mathbb{N}$ and $\mathbf{c}(N+1) = \mathbf{c}(1)$, where N is the number of samples . From this contour, different signatures can be calculated:

- a complex signature:

$$\mathbf{s}(t) = \mathbf{x}(t) + i\mathbf{y}(t)$$

  where $i^2 = -1$
- the centroid distance signature:

$$\mathbf{s}(t) = \sqrt{(\mathbf{x}(t) - x_c)^2 + (\mathbf{y}(t) - y_c)^2}$$

  where (x_c, y_c) is the centroid of $\mathbf{c}$
- a direction signature, i.e. direction of the tangent line at a point t:

$$\mathbf{s}(t) = \cos^{-1}\left(\frac{\mathbf{y}(t+1) - \mathbf{y}(t-1)}{\|\mathbf{c}(t+1) - \mathbf{c}(t-1)\|}\right)$$

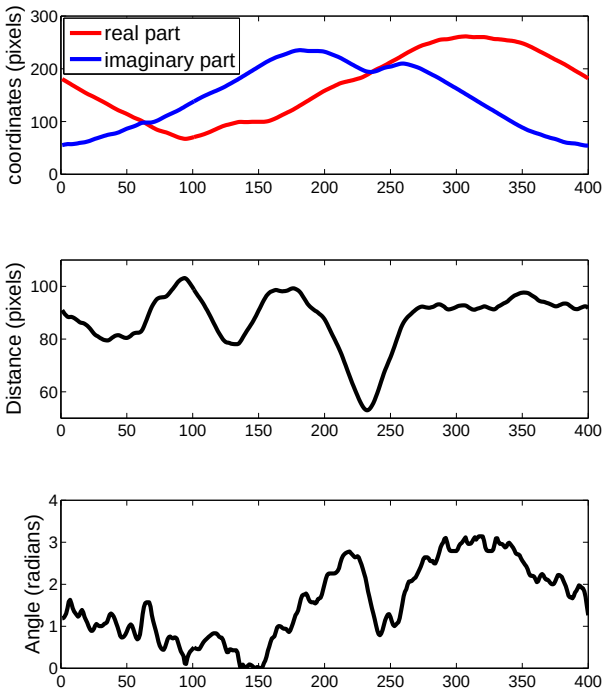


Figure 2: An example of different signatures. These are the signatures for the contour shown in Fig. 1.e. From top to bottom the complex, the centroid distance and the direction signature are shown

For a more detailed overview of signature functions we refer to [6, 13]. In Fig. 2 an example is shown of the defined signatures. This are the signatures for the nucleus shown in Fig. 1.e. These signatures are from top to bottom: complex, centroid distance and the direction signature.

3.2 Signature Matching

Note that each point on a signature exactly corresponds to one point on the curve. So matching contour points is equivalent to matching points on a signature. For the matching of two signatures, we propose a matching based on *Dynamic Time Warping* or *dog-man distance*. The DTW distance be-

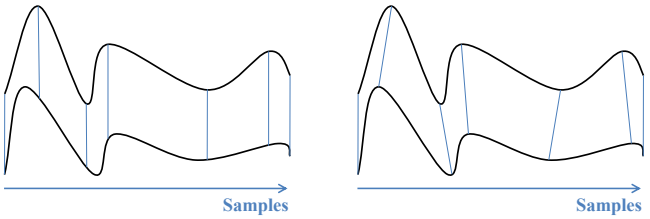


Figure 3: An example of matching two curves. On the left a naive matching where the matching is done using the parameterization. On the right the matching with the minimal cumulative difference between the samples. It is this matching which will correspond with the warping used for DTW.

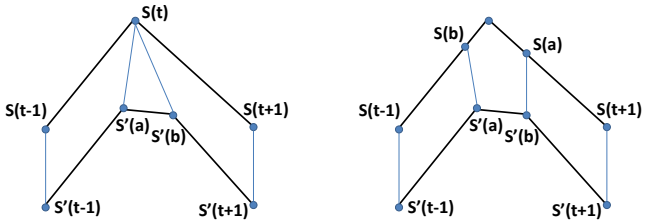


Figure 4: An example of resampling the curves based on the warping function of DTW

tween two signals is analogous of a man and a dog each waking with a different path, i.e. the signals. The dog can walk on a different speed by giving it a longer leash. Fig. 3 shows on the left two paths, where both paths are scanned at the same speed, so man and dog are walking on the same speed. On the right of the figure, different speeds result in a a better match, i.e. the cumulative difference between these samples is less. The dog-man distance is then the shortest possible cumulative difference between both paths, by changing the scanning speed, i.e.

$$d_{DTW}(\mathbf{s1}(.),\mathbf{s2}(.)) = \min_{\gamma} \sum_{t} \|s1(t) - s2(\gamma(t))\| \quad (3)$$

where $\gamma(.)$ is a warping function, i.e. any monotonic function mapping $[1,N]$ on $[1,N]$. This can be calculated using dynamic programming:

$$D(i,j) = d(i,j) + \min(\alpha D(i,j-1), \alpha D(i-1,j), D(i-1,j-1)) \quad (4)$$

with $d(i,j) = \|\mathbf{s1}(i) - \mathbf{s2}(j)\|$ and α a real number greater or equal then one. If α equal to one, all warping functions are considered equally good. In order to penalize warping functions where to many points of one signal are matched to a single point in the other signal, one can set $\alpha > 1$.

The distance itself is of little importance to us, but the warping function resulting in the minimal distance is, i.e. $\gamma(.)$. Using linear interpolation the signals can be considered continuous, based on the warping function the signals will be resampled in such a way that there is a unique point to point correspondence. Consider the left case of Fig. 4, where one point, $s(t)$ of the signal is matched with multiple points in the other signal. First define $s(t-1)$ and $s(t+1)$ the last predecessor and first successor of $s(t)$ with a unique match, e.g. $s'(t-1)$ and $s'(t+1)$ respectively. The point $s(t)$ will be replaced by the number of points between $s(t-1)$ and $s(t+1)$: for each point $s'(i)$ between $s'(t-1)$ and $s'(t+1)$ a new sample is added on $s(.)$. These new samples are calculated in such a way that

$$\frac{al(s'(.),t-1,i)}{al(s'(.),t-1,t+1)} = \frac{al(s(.),t-1,i)}{al(s(.),t-1,t+1)} \quad (5)$$

where $al(k(.),l,m)$ is the arc length from point l to m over the curve $k(.)$.

4. CELL MAPPING

Based on the previous contour matching we get for each point on the contour at frame t a unique point on the contour at frame $(t+1)$. So for each point on the contour we assume to know the exact position in the next frame. For the mapping of points not on the contour, $P = (x,y)$, we propose interpolation with biharmonic splines:

$$P_{t+1} = a + a_x x + a_y y + \sum_{j=1}^{N} w_j \|P_t - \mathbf{c}(j)\| \quad (6)$$

where a, a_x, a_y and w_j are a set of weighting coefficients. These coefficients are calculated in such a way that:

- points on the contour are mapped to the corresponding points according to the contour matching
- the weighting vector $\mathbf{w}$ is orthogonal to $\mathbf{x}$ and $\mathbf{y}$, i.e.

$$\sum_{i=1}^{N} w_i x_i = \sum_{i=1}^{N} w_i y_i = 0 \quad (7)$$

These weighting coefficients can be found by solving a linear system of equations, which can be done in a fast way as proposed by Beatson et al. [1].

5. RESULTS

Two datasets where used for the validation of the proposed method. Both datasets contain an isolated fibroblast cell nucleus of a patient with a homozygous nonsense mutation in lamin A/C gene. These are cells where the nucleus can have big deformations [4]. The datasets where captured using confocal controlled light-exposure microscopy [3], resulting in images of 288×288 pixels. These cells were transfected with a construct expressing a fluorescent fusion protein targeting the telomeres (TRF2-mCitrine). Non-bound protein diffuses freely throughout the nucleus and allows for demarcating the boundaries for contour identification. For both validation sets the telomeres where manually tracked by the author using MTrackJ[1]. The centroids of these tracked telomeres will serve as ground truth for the validation. The telomeres them self might move locally, but this small local motion, e.g. 1 to 2 pixels, is neglectable compared to the big motion introduced by deforming nuclei.

In Fig. 5 three examples of the mapping are shown. The red and blue curves are the contours of a nucleus in subsequent time frames. The red and blue dots are the centroids

[1] http://www.imagescience.org/meijering/software/mtrackj/

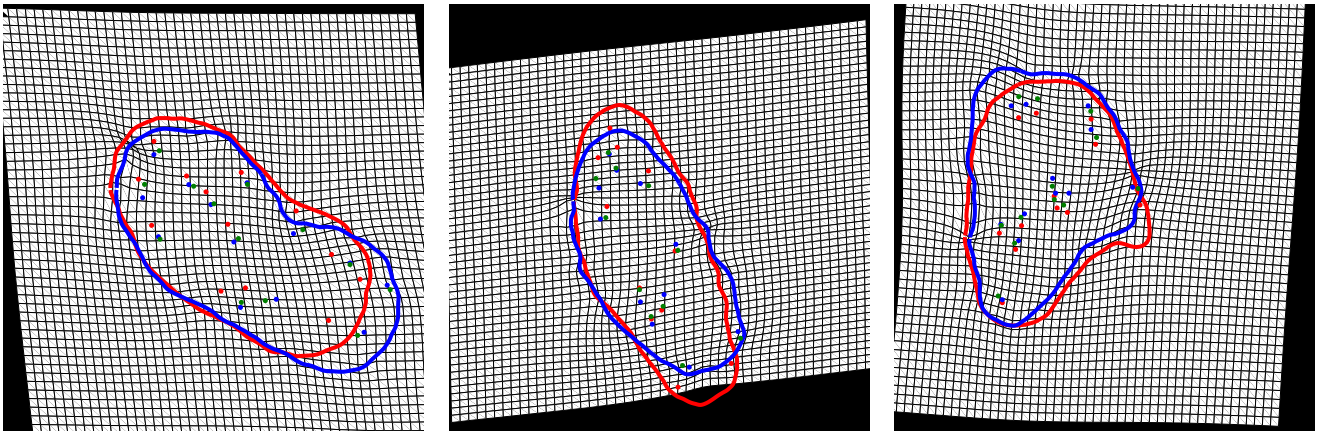


Figure 5: Three examples of mapping telomeres from a frame to a subsequent time frame. The red and blue curves represent the nucleus contour in the concerning frames. The he red and blue dots are the telomere centroids in the frame corresponding to respectively the red and blue curve. The green dots are the mapping of the red dots based on the warping of the red and blue contour.

of the telomeres corresponding to respectively the red and blue contours. The green dots are the location of the red dots mapped with the proposed technique using the centroid distance signatures. So the green dots are the prediction of the location of the blue dots. As can be seen does the mapping closely approximate the real telomere location in the next time frame.

In Table 1 the results are shown for the full datasets. The first column shows the dataset. In the first dataset 12 telomeres where tracked during 12 frames. The second dataset consists of 11 telomeres which were tracked for 11 frames In the second column the used signature is shown. Distance means the non processed data, i.e. the centroid coordinates are used as they are extracted out of the images, without any mapping. The third column shows the average Euclidean distance between the telomeres in the one frame and where the telomeres are ought to be according to the mapping of the location of the telomere in the previous frame. This distance or error is expressed in pixels. In the next column the median error is shown, also in pixels. The fifth column shows the variance of the error. For the sixth column the telomeres of a frame where matched with the telomeres in the previous frame using the *Iterative Closest Point (ICP)* method. The amount of false matches for the full sequence are mentioned in this column. The last columns give measures of the improvement of the proposed techniques compares to using the unprocessed data, i.e. the "Distance" row. Both the average as the median improvements are shown. And for both measurements this is expressed both absolute, i.e. the distance in pixels that the proposed method works better, as relative, i.e. the ratio of the error using the proposed method over the error using the non processed data.

The proposed method results in a significant improvement compared to not mapping the nuclei as can be seen in the table, although the amount of improvement differs for both datasets. The centroid distance and direction signature perform best for both datasets. Since the absolute error is similar for both signatures, the direction signature is slightly in favor based on the ICP results.

6. CONCLUSION

In this paper a new mapping algorithm is proposed for the mapping of cell nuclei. This mapping does not impose any shape constraints, which allows it to map cell nuclei of pathological cells. The proposed technique works in two steps. First the nucleus contours are matched. Three different matching techniques are defined and tested. Based on this contour matching the full nucleus is mapped using biharmonic interpolation. The proposed method was tested on two real datasets, which both show significant improvement.

7. ACKNOWLEDGMENT

Jonas De Vylder is funded by the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT).

REFERENCES

[1] R. K. Beatson, M. J. D. Powell, and A. M. Tan. Fast evaluation of polyharmonic splines in three dimensions. *IMA Journal of Numerical Analysis*, 27:427–450, 2007.

[2] X. W. Chen, X. B. Zhou, and S. T. C. Wong. Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. *IEEE Transactions on Biomedical Engineering*, 53(4):762–766, 2006.

[3] W. H. De Vos, R. A. Hoebe, G. H. Joss, W. Haffinans, S. Baatout, P. Van Oostveldt, and E. M. M. Manders. Controlled light exposure microscopy reveals dynamic telomere microterritories throughout the cell cycle. *Cytometry Part A*, 75A(5):428–439, 2009.

[4] W. H. De Vos, F. Houben, R. A. Hoebe, R. Hennekam, B. van Engelen, E. M. Manders, F. C. Ramaekers, J. L. Broers, and V. O. P. Increased plasticity of the nuclear envelope and hypermobility of telomeres due to the loss of a-type lamins. *Biochimica et Biophysica Acta (BBA) - General Subjects*, In Press, Corrected Proof:–, 2010.

[5] E. Gladilin, S. Goetze, J. Mateos-Langerak, R. Van Driel, R. Eils, and K. Rohr. Shape normalization of 3d cell nuclei using elastic spherical mapping. *Journal of Microscopy-Oxford*, 231(1):105–114, 2008.

data set	method	av. err.	med. err.	σ^2	ICP err.	av. improvement		med. improvement	
						absolute	relative	absolute	relative
nucleus 1	Distance	7.7	7.2	12	15				
	Complex	3.8	3.1	7.4	4	3.8	53%	3.3	49%
	Centroid dist.	2.5	2.0	2.6	3	5.2	48%	4.9	28%
	Direction	3.0	2.7	4.0	0	4.7	50%	4.0	35%
nucleus 2	Distance	8.2	7.7	11.5	30				
	Complex	6.8	6.9	10.4	23	1.5	89%	1.5	86%
	Centroid dist.	6.3	5	17.3	16	2	87%	2.3	69%
	Direction	5.2	4.7	7.5	8	3.1	67%	2.8	65%

Table 1: Error measurements of the proposed methods

[6] V. V. Kindratenko. On using functions to describe the shape. *Journal of Mathematical Imaging and Vision*, 18(3):225–245, 2003.

[7] G. Li, T. Liu, J. Nie, L. Guo, J. Chen, J. Zhu, W. Xia, A. Mara, S. Holley, and S. T. C. Wong. Segmentation of touching cell nuclei using gradient flow tracking. *Journal of Microscopy-Oxford*, 231(1):47–58, 2008.

[8] S. F. Louis, B. J. Vermolen, Y. Garini, I. T. Young, A. Guffei, Z. Lichtensztejn, F. Kuttler, T. C. Y. Chuang, S. Moshir, V. Mougey, A. Y. C. Chuang, P. D. Kerr, T. Fest, P. Boukamp, and S. Mai. c-myc induces chromosomal rearrangements through telomere and chromosome remodeling in the interphase nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27):9613–9618, 2005.

[9] S. G. Megason and S. E. Fraser. Digitizing life at the level of the cell: high-performance laser-scanning microscopy and image analysis for in toto imaging of development. *Mechanisms of Development*, 120(11):1407–1420, 2003.

[10] N. Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on Systems, Man and Cybernetics*, 9:62–66, 1976.

[11] V. Raz, B. J. Vermolen, Y. Garini, J. J. M. Onderwater, M. A. Mommaas-Kienhuis, A. J. Koster, I. T. Young, H. Tanke, and R. W. Dirks. The nuclear lamina promotes telomere aggregation and centromere peripheral localization during senescence of human mesenchymal stem cells. *Journal of Cell Science*, 121(24):4018–4028, 2008.

[12] P. K. Yan, X. B. Zhou, M. Shah, and S. T. C. Wong. Automatic segmentation of high-throughput rnai fluorescent cellular images. *IEEE Transactions on Information Technology in Biomedicine*, 12(1):109–117, 2008.

[13] D. S. Zhang and G. J. Lu. Review of shape representation and description techniques. *Pattern Recognition*, 37(1):1–19, 2004.